

**SYSTEM AND METHOD FOR COMBINING VOICE  
ANNOTATION AND RECOGNITION SEARCH CRITERIA WITH  
TRADITIONAL SEARCH CRITERIA INTO METADATA**

Michelle Lehmeier  
2027 Wimbledon Drive  
Loveland, Colorado 80538  
Citizenship: United States

Edward Beeman  
33766 Cliff Road  
Windsor, Colorado 80550  
Citizenship: United States

Robert Sobol  
2313 Idledale Drive  
Ft. Collins, Colorado 80526  
Citizenship: United States

**TECHNICAL FIELD**

The present invention relates to data access, and more particularly, to the identification and retrieval of specific data through the use of keywords in textual documents or key objects in image documents and by verbal descriptions of the data contained in all types of documents.

## BACKGROUND

The generation and use of keywords to index, store and retrieve textual documents is well known in the prior art. These keywords are typically generated by the document's creator and are used as an indication of document content and aids in the selection and retrieval of applicable documents from a document or image database. Additionally, it is well known in the prior art that the body of textual documents can be searched for specific words or phrases to find a textual document, or an area of the document which is of interest to the searcher. Similarly, computer directories or subdirectories may be searched to identify documents which pertain to certain subjects, areas of interest, or topics. Keywords can also be associated with these subdirectories, using a subdirectory naming convention, to indicate the data which is contained within a subdirectory. While various search engines provide for searching of written documents, searching of other forms of materials, such as images, is not well supported. Further, most documents and other databases do not readily support searching other than in the context of the object stored and, more commonly, by file name or other text based searching routines.

## SUMMARY OF THE INVENTION

The present invention is directed to a system and method which provides for enhanced indexing, categorization, and retrieval of documents by, according to one aspect of the invention, combining index terms derived from document content and file information with user provided information, such as spoken commentary. The spoken commentary may be stored as a digitized audio file and/or subjected to processing, such as speech recognition, converting the spoken commentary to, for example, text. The text may then be parsed (searched and portions extracted for use) to identify and extract additional searchable terms and phrases and/or be used to otherwise enhance and support document access, search, identification, and retrieval capabilities.

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80  
85  
90  
95  
100  
105  
110  
115  
120  
125  
130  
135  
140  
145  
150  
155  
160  
165  
170  
175  
180  
185  
190  
195  
200  
205  
210  
215  
220  
225  
230  
235  
240  
245  
250  
255  
260  
265  
270  
275  
280  
285  
290  
295  
300  
305  
310  
315  
320  
325  
330  
335  
340  
345  
350  
355  
360  
365  
370  
375  
380  
385  
390  
395  
400  
405  
410  
415  
420  
425  
430  
435  
440  
445  
450  
455  
460  
465  
470  
475  
480  
485  
490  
495  
500  
505  
510  
515  
520  
525  
530  
535  
540  
545  
550  
555  
560  
565  
570  
575  
580  
585  
590  
595  
600  
605  
610  
615  
620  
625  
630  
635  
640  
645  
650  
655  
660  
665  
670  
675  
680  
685  
690  
695  
700  
705  
710  
715  
720  
725  
730  
735  
740  
745  
750  
755  
760  
765  
770  
775  
780  
785  
790  
795  
800  
805  
810  
815  
820  
825  
830  
835  
840  
845  
850  
855  
860  
865  
870  
875  
880  
885  
890  
895  
900  
905  
910  
915  
920  
925  
930  
935  
940  
945  
950  
955  
960  
965  
970  
975  
980  
985  
990  
995  
1000  
1005  
1010  
1015  
1020  
1025  
1030  
1035  
1040  
1045  
1050  
1055  
1060  
1065  
1070  
1075  
1080  
1085  
1090  
1095  
1100  
1105  
1110  
1115  
1120  
1125  
1130  
1135  
1140  
1145  
1150  
1155  
1160  
1165  
1170  
1175  
1180  
1185  
1190  
1195  
1200  
1205  
1210  
1215  
1220  
1225  
1230  
1235  
1240  
1245  
1250  
1255  
1260  
1265  
1270  
1275  
1280  
1285  
1290  
1295  
1300  
1305  
1310  
1315  
1320  
1325  
1330  
1335  
1340  
1345  
1350  
1355  
1360  
1365  
1370  
1375  
1380  
1385  
1390  
1395  
1400  
1405  
1410  
1415  
1420  
1425  
1430  
1435  
1440  
1445  
1450  
1455  
1460  
1465  
1470  
1475  
1480  
1485  
1490  
1495  
1500  
1505  
1510  
1515  
1520  
1525  
1530  
1535  
1540  
1545  
1550  
1555  
1560  
1565  
1570  
1575  
1580  
1585  
1590  
1595  
1600  
1605  
1610  
1615  
1620  
1625  
1630  
1635  
1640  
1645  
1650  
1655  
1660  
1665  
1670  
1675  
1680  
1685  
1690  
1695  
1700  
1705  
1710  
1715  
1720  
1725  
1730  
1735  
1740  
1745  
1750  
1755  
1760  
1765  
1770  
1775  
1780  
1785  
1790  
1795  
1800  
1805  
1810  
1815  
1820  
1825  
1830  
1835  
1840  
1845  
1850  
1855  
1860  
1865  
1870  
1875  
1880  
1885  
1890  
1895  
1900  
1905  
1910  
1915  
1920  
1925  
1930  
1935  
1940  
1945  
1950  
1955  
1960  
1965  
1970  
1975  
1980  
1985  
1990  
1995  
2000  
2005  
2010  
2015  
2020  
2025  
2030  
2035  
2040  
2045  
2050  
2055  
2060  
2065  
2070  
2075  
2080  
2085  
2090  
2095  
2100  
2105  
2110  
2115  
2120  
2125  
2130  
2135  
2140  
2145  
2150  
2155  
2160  
2165  
2170  
2175  
2180  
2185  
2190  
2195  
2200  
2205  
2210  
2215  
2220  
2225  
2230  
2235  
2240  
2245  
2250  
2255  
2260  
2265  
2270  
2275  
2280  
2285  
2290  
2295  
2300  
2305  
2310  
2315  
2320  
2325  
2330  
2335  
2340  
2345  
2350  
2355  
2360  
2365  
2370  
2375  
2380  
2385  
2390  
2395  
2400  
2405  
2410  
2415  
2420  
2425  
2430  
2435  
2440  
2445  
2450  
2455  
2460  
2465  
2470  
2475  
2480  
2485  
2490  
2495  
2500  
2505  
2510  
2515  
2520  
2525  
2530  
2535  
2540  
2545  
2550  
2555  
2560  
2565  
2570  
2575  
2580  
2585  
2590  
2595  
2600  
2605  
2610  
2615  
2620  
2625  
2630  
2635  
2640  
2645  
2650  
2655  
2660  
2665  
2670  
2675  
2680  
2685  
2690  
2695  
2700  
2705  
2710  
2715  
2720  
2725  
2730  
2735  
2740  
2745  
2750  
2755  
2760  
2765  
2770  
2775  
2780  
2785  
2790  
2795  
2800  
2805  
2810  
2815  
2820  
2825  
2830  
2835  
2840  
2845  
2850  
2855  
2860  
2865  
2870  
2875  
2880  
2885  
2890  
2895  
2900  
2905  
2910  
2915  
2920  
2925  
2930  
2935  
2940  
2945  
2950  
2955  
2960  
2965  
2970  
2975  
2980  
2985  
2990  
2995  
3000  
3005  
3010  
3015  
3020  
3025  
3030  
3035  
3040  
3045  
3050  
3055  
3060  
3065  
3070  
3075  
3080  
3085  
3090  
3095  
3100  
3105  
3110  
3115  
3120  
3125  
3130  
3135  
3140  
3145  
3150  
3155  
3160  
3165  
3170  
3175  
3180  
3185  
3190  
3195  
3200  
3205  
3210  
3215  
3220  
3225  
3230  
3235  
3240  
3245  
3250  
3255  
3260  
3265  
3270  
3275  
3280  
3285  
3290  
3295  
3300  
3305  
3310  
3315  
3320  
3325  
3330  
3335  
3340  
3345  
3350  
3355  
3360  
3365  
3370  
3375  
3380  
3385  
3390  
3395  
3400  
3405  
3410  
3415  
3420  
3425  
3430  
3435  
3440  
3445  
3450  
3455  
3460  
3465  
3470  
3475  
3480  
3485  
3490  
3495  
3500  
3505  
3510  
3515  
3520  
3525  
3530  
3535  
3540  
3545  
3550  
3555  
3560  
3565  
3570  
3575  
3580  
3585  
3590  
3595  
3600  
3605  
3610  
3615  
3620  
3625  
3630  
3635  
3640  
3645  
3650  
3655  
3660  
3665  
3670  
3675  
3680  
3685  
3690  
3695  
3700  
3705  
3710  
3715  
3720  
3725  
3730  
3735  
3740  
3745  
3750  
3755  
3760  
3765  
3770  
3775  
3780  
3785  
3790  
3795  
3800  
3805  
3810  
3815  
3820  
3825  
3830  
3835  
3840  
3845  
3850  
3855  
3860  
3865  
3870  
3875  
3880  
3885  
3890  
3895  
3900  
3905  
3910  
3915  
3920  
3925  
3930  
3935  
3940  
3945  
3950  
3955  
3960  
3965  
3970  
3975  
3980  
3985  
3990  
3995  
4000  
4005  
4010  
4015  
4020  
4025  
4030  
4035  
4040  
4045  
4050  
4055  
4060  
4065  
4070  
4075  
4080  
4085  
4090  
4095  
4100  
4105  
4110  
4115  
4120  
4125  
4130  
4135  
4140  
4145  
4150  
4155  
4160  
4165  
4170  
4175  
4180  
4185  
4190  
4195  
4200  
4205  
4210  
4215  
4220  
4225  
4230  
4235  
4240  
4245  
4250  
4255  
4260  
4265  
4270  
4275  
4280  
4285  
4290  
4295  
4300  
4305  
4310  
4315  
4320  
4325  
4330  
4335  
4340  
4345  
4350  
4355  
4360  
4365  
4370  
4375  
4380  
4385  
4390  
4395  
4400  
4405  
4410  
4415  
4420  
4425  
4430  
4435  
4440  
4445  
4450  
4455  
4460  
4465  
4470  
4475  
4480  
4485  
4490  
4495  
4500  
4505  
4510  
4515  
4520  
4525  
4530  
4535  
4540  
4545  
4550  
4555  
4560  
4565  
4570  
4575  
4580  
4585  
4590  
4595  
4600  
4605  
4610  
4615  
4620  
4625  
4630  
4635  
4640  
4645  
4650  
4655  
4660  
4665  
4670  
4675  
4680  
4685  
4690  
4695  
4700  
4705  
4710  
4715  
4720  
4725  
4730  
4735  
4740  
4745  
4750  
4755  
4760  
4765  
4770  
4775  
4780  
4785  
4790  
4795  
4800  
4805  
4810  
4815  
4820  
4825  
4830  
4835  
4840  
4845  
4850  
4855  
4860  
4865  
4870  
4875  
4880  
4885  
4890  
4895  
4900  
4905  
4910  
4915  
4920  
4925  
4930  
4935  
4940  
4945  
4950  
4955  
4960  
4965  
4970  
4975  
4980  
4985  
4990  
4995  
5000  
5005  
5010  
5015  
5020  
5025  
5030  
5035  
5040  
5045  
5050  
5055  
5060  
5065  
5070  
5075  
5080  
5085  
5090  
5095  
5100  
5105  
5110  
5115  
5120  
5125  
5130  
5135  
5140  
5145  
5150  
5155  
5160  
5165  
5170  
5175  
5180  
5185  
5190  
5195  
5200  
5205  
5210  
5215  
5220  
5225  
5230  
5235  
5240  
5245  
5250  
5255  
5260  
5265  
5270  
5275  
5280  
5285  
5290  
5295  
5300  
5305  
5310  
5315  
5320  
5325  
5330  
5335  
5340  
5345  
5350  
5355  
5360  
5365  
5370  
5375  
5380  
5385  
5390  
5395  
5400  
5405  
5410  
5415  
5420  
5425  
5430  
5435  
5440  
5445  
5450  
5455  
5460  
5465  
5470  
5475  
5480  
5485  
5490  
5495  
5500  
5505  
5510  
5515  
5520  
5525  
5530  
5535  
5540  
5545  
5550  
5555  
5560  
5565  
5570  
5575  
5580  
5585  
5590  
5595  
5600  
5605  
5610  
5615  
5620  
5625  
5630  
5635  
5640  
5645  
5650  
5655  
5660  
5665  
5670  
5675  
5680  
5685  
5690  
5695  
5700  
5705  
5710  
5715  
5720  
5725  
5730  
5735  
5740  
5745  
5750  
5755  
5760  
5765  
5770  
5775  
5780  
5785  
5790  
5795  
5800  
5805  
5810  
5815  
5820  
5825  
5830  
5835  
5840  
5845  
5850  
5855  
5860  
5865  
5870  
5875  
5880  
5885  
5890  
5895  
5900  
5905  
5910  
5915  
5920  
5925  
5930  
5935  
5940  
5945  
5950  
5955  
5960  
5965  
5970  
5975  
5980  
5985  
5990  
5995  
6000  
6005  
6010  
6015  
6020  
6025  
6030  
6035  
6040  
6045  
6050  
6055  
6060  
6065  
6070  
6075  
6080  
6085  
6090  
6095  
6100  
6105  
6110  
6115  
6120  
6125  
6130  
6135  
6140  
6145  
6150  
6155  
6160  
6165  
6170  
6175  
6180  
6185  
6190  
6195  
6200  
6205  
6210  
6215  
6220  
6225  
6230  
6235  
6240  
6245  
6250  
6255  
6260  
6265  
6270  
6275  
6280  
6285  
6290  
6295  
6300  
6305  
6310  
6315  
6320  
6325  
6330  
6335  
6340  
6345  
6350  
6355  
6360  
6365  
6370  
6375  
6380  
6385  
6390  
6395  
6400  
6405  
6410  
6415  
6420  
6425  
6430  
6435  
6440  
6445  
6450  
6455  
6460  
6465  
6470  
6475  
6480  
6485  
6490  
6495  
6500  
6505  
6510  
6515  
6520  
6525  
6530  
6535  
6540  
6545  
6550  
6555  
6560  
6565  
6570  
6575  
6580  
6585  
6590  
6595  
6600  
6605  
6610  
6615  
6620  
6625  
6630  
6635  
6640  
6645  
6650  
6655  
6660  
6665  
6670  
6675  
6680  
6685  
6690  
6695  
6700  
6705  
6710  
6715  
6720  
6725  
6730  
6735  
6740  
6745  
6750  
6755  
6760  
6765  
6770  
6775  
6780  
6785  
6790  
6795  
6800  
6805  
6810  
6815  
6820  
6825  
6830  
6835  
6840  
6845  
6850  
6855  
6860  
6865  
6870  
6875  
6880  
6885  
6890  
6895  
6900  
6905  
6910  
6915  
6920  
6925  
6930  
6935  
6940  
6945  
6950  
6955  
6960  
6965  
6970  
6975  
6980  
6985  
6990  
6995  
7000  
7005  
7010  
7015  
7020  
7025  
7030  
7035  
7040  
7045  
7050  
7055  
7060  
7065  
7070  
7075  
7080  
7085  
7090  
7095  
7100  
7105  
7110  
7115  
7120  
7125  
7130  
7135  
7140  
7145  
7150  
7155  
7160  
7165  
7170  
7175  
7180  
7185  
7190  
7195  
7200  
7205  
7210  
7215  
7220  
7225  
7230  
7235  
7240  
7245  
7250  
7255  
7260  
7265  
7270  
7275  
7280  
7285  
7290  
7295  
7300  
7305  
7310  
7315  
7320  
7325  
7330  
7335  
7340  
7345  
7350  
7355  
7360  
7365  
7370  
7375  
7380  
7385  
7390  
7395  
7400  
7405  
7410  
7415  
7420  
7425  
7430  
7435  
7440  
7445  
7450  
7455  
7460  
7465  
7470  
7475  
7480  
7485  
7490  
7495  
7500  
7505  
7510  
7515  
7520  
7525  
7530  
7535  
7540  
7545  
7550  
7555  
7560  
7565  
7570  
7575  
7580  
7585  
7590  
7595  
7600  
7605  
7610  
7615  
7620  
7625  
7630  
7635  
7640  
7645  
7650  
7655  
7660  
7665  
7670  
7675  
7680  
7685  
7690  
7695  
7700  
7705  
7710  
7715  
7720  
7725  
7730  
7735  
7740  
7745  
7750  
7755  
7760  
7765  
7770  
7775  
7780  
7785  
7790  
7795  
7800  
7805  
7810  
7815  
7820  
7825  
7830  
7835  
7840  
7845  
7850  
7855  
7860  
7865  
7870  
7875  
7880  
7885  
7890  
7895  
7900  
7905  
7910  
7915  
7920  
7925  
7930  
7935  
7940  
7945  
7950  
7955  
7960  
7965  
7970  
7975  
7980  
7985  
7990  
7995  
8000  
8005  
8010  
8015  
8020  
8025  
8030  
8035  
8040  
8045  
8050  
8055  
8060  
8065  
8070  
8075  
8080  
8085  
8090  
8095  
8100  
8105  
8110  
8115  
8120  
8125  
8130  
8135  
8140  
8145  
8150  
8155  
8160  
8165  
8170  
8175  
8180  
8185  
8190  
8195  
8200  
8205  
8210  
8215  
8220  
8225  
8230  
8235  
8240  
8245  
8250  
8255  
8260  
8265  
8270  
8275  
8280  
8285  
8290  
8295  
8300  
8305  
8310  
8315  
8320  
8325  
8330  
8335  
8340  
8345  
8350  
8355  
8360  
8365  
8370  
8375  
8380  
8385  
8390  
8395  
8400  
8405  
8410  
8415  
8420  
8425  
8430  
8435  
8440  
8445  
8450  
8455  
8460  
8465  
8470  
8475  
8480  
8485  
8490  
8495  
8500  
8505  
8510  
8515  
8520  
8525  
8530  
8535  
8540  
8545  
8550  
8555  
8560  
8565  
8570  
8575  
8580  
8585  
8590  
8595  
8600  
8605  
8610  
8615  
8620  
8625  
8630  
8635  
8640  
8645  
8650  
8655  
8660  
8665  
8670  
8675  
8680  
8685  
8690  
8695  
8700  
8705  
8710  
8715  
8720  
8725  
8730  
8735  
8740  
8745  
8750  
8755  
8760  
8765  
8770  
8775  
8780  
8785  
8790  
8795  
8800  
8805  
8810  
8815  
8820  
8825  
8830  
8835  
8840  
8845  
8850  
8855  
8860  
8865  
8870  
8875  
8880  
8885  
8890  
8895  
8900  
8905  
8910  
8915  
8920  
8925  
8930  
8935  
8940  
8945  
8950  
8955  
8960  
8965  
8970  
8975  
8980  
8985  
8990  
8995  
9000  
9005  
9010  
9015  
9020  
9025  
9030  
9035  
9040  
9045  
9050  
9055  
9060  
9065  
9070  
9075  
9080  
9085  
9090  
9095  
9100  
9105  
9110  
9115  
9120  
9125  
9130  
9135  
9140  
9145  
9150  
9155  
9160  
9165  
9170  
9175  
9180  
9185  
9190  
9195  
9200  
9205  
9210  
9215  
9220  
9225  
9230  
9235  
9240  
9245  
9250  
9255  
9260  
9265  
9270  
9275  
9280  
9285  
9290  
9295  
9300  
9305  
9310  
9315  
9320  
9325  
9330  
9335  
9340  
9345  
9350  
9355  
9360  
9365  
9370  
9375  
9380  
9385  
9390  
9395  
9400  
9405  
9410  
9415  
9420  
9425  
9430  
9435  
9440  
9445  
9450  
9455  
9460  
9465  
9470  
9475  
9480  
9485  
9490  
9495  
9500  
9505  
9510  
9515  
9520  
9525  
9530  
9535  
9540  
9545  
9550  
9555  
9560  
9565  
9570  
9575  
9580  
9585  
9590  
9595  
9600  
9605  
9610  
9615  
9620  
9625  
9630  
9635  
9640  
9645  
9650  
9655  
9660  
9665  
9670  
9675  
9680  
9685  
9690  
9695  
9700  
9705  
9710  
9715  
9720  
9725  
9730  
9735  
9740  
9745  
9750  
9755  
9760  
9765  
9770  
9775  
9780  
9785  
9790  
9795  
9800  
9805  
9810  
9815  
9820  
9825  
9830  
9835  
9840  
9845  
9850  
9855  
9860  
9865  
9870  
9875  
9880  
9885  
9890  
9895  
9900  
9905  
9910  
9915  
9920  
9925  
9930  
9935  
9940  
9945  
9950  
9955  
9960  
9965  
9970  
9975  
9980  
9985  
9990  
9995  
10000  
10005  
10010  
10015  
10020  
10025  
10030  
10035  
10040  
10045  
10050  
10055  
10060  
10065  
10070  
10075  
10080  
10085  
10090  
10095  
10100  
10105  
10110  
10115  
10120  
10125  
10130  
10135  
10140  
10145  
10150  
10155  
10160  
10165  
10170  
10175  
10180  
10185  
1019

Another embodiment of the present invention includes a system for storing documents in an electronic storage media including a means for obtaining data tags pertaining to certain characteristics of each document which are selected from a list of recognized characters, semantics processing, object and voice recognition and a means for associating the data with the document.

5

09873687 060401  
1010910 289E2860

BRIEF DESCRIPTION OF THE DRAWING

FIGURE 1 is a block diagram of a method of differentiating textual documents;

FIGURE 2 is a block diagram of a method of differentiating image or picture documents;

FIGURE 3 is a block diagram showing the use of voice annotation and recognition in conjunction with additional search criteria;

FIGURE 4 is an example of a database which associates documents with their keywords, keynotes and key objects; and

FIGURE 5 is a block diagram of a system which implements the current invention.

09873687-060401

## DETAILED DESCRIPTION

The present invention is directed to a system such as a document retrieval system, and a methodology for identifying documents which can be applied to both textual documents as well as photographic documents or images. The invention is equally applicable to an image storing system for storing document in electronic media. Typically, document users identify  
5 desired documents by the file name or through keyword searches of computer text files.

When many similar documents are stored, differentiating the various documents by file name in a meaningful way becomes difficult, if not impossible. The next step in differentiating documents is to supply document keywords or other groupings to indicate the information the documents contain or that are otherwise associated with the document (*e.g.*, synonyms of terminology used in document, related concepts, etc.). These words or groupings can consist of keywords, or sentences which describe the information contained in the textual document. Similarly, images stored on electronic media can be differentiated from one another by the image's file name. These images can be further differentiated by their placement within the electronic media. For instance, distinct media or separate subdirectories within a media can be created which include only images of a certain subject matter. Thus, for example, if a user stores all their photographs on electronic media, a single diskette can be dedicated to vacation pictures from 1995, a separate diskette can be dedicated to vacation pictures from 1996, and a third diskette can be dedicated to pictures from the vacation in 1997. These storage techniques mimic traditional photo albums. Alternatively, subdirectories can be used on a  
20 single recording device (*e.g.*, hard drive) to differentiate photographs from various time periods or vacations. The current invention builds and expands on these capabilities by allowing the user to associate spoken words, phrases or text extracted from an image with annotation on the textual or image document to identify documents, access the document, or differentiate the document among other unrelated documents.

25 One object of this invention is to combine keyword capability with other user-supplied information to identify and access textual documents. Additionally, a further object of the invention is to enable images stored by a computer to be indexed, sorted, and accessed by reference to objects included within the images and/or by user-supplied information. A

still further object of the invention is a method in which an individual may annotate a document and use the annotation to search and retrieve documents and other objects from a database.

Referring now to FIGURE 1, a procedure for differentiating textual documents is illustrated. Textual documents can be the result of word processed documents or scanned documents. Scanned documents, for instance, are created from optical scanning hard copies of documents into a file. These scanned documents are fed to a character recognition program (*i.e.*, an Optical Character Recognition (OCR) program), which translates pixel image information contained within the scanned document to textual information. This function is performed by character recognition block 101 of FIGURE 1. For textual documents generated by a word processing program this step may be omitted. The resulting textual information can then be accessed by a word processing program to delete, change, or add to the information contained within the body of the textual document. The textual document can also be accessed by semantics processing block 102 to identify keywords associated within the textual information. Such semantics processing programs may respond to the number of times a specific word appears within the document, the keywords assigned to the textual document by the user, or by any other method which distills the textual information down to a number of keywords which describe and/or characterize the textual document. These keywords can then be processed by metadata program 103 which will assign the keywords as indexes to the associated textual document. This assignment, may for example take the form of a table which associates keywords with file or document names. FIGURE 4 depicts one representation of this association. This metadata may take several different forms, including a database which tracks document names or file names with their associated keywords.

Referring now to FIGURE 2, keywords can also be associated with images, or digital pictures as shown in process 200. Digital pictures, or scanned images can be processed by object recognition program 201 to identify the specific objects included within the digital photograph or scanned image. Object recognition program 201 may consist of software which detects edges between various objects within a digital photograph or scanned image,

and may identify the images contained within the picture or scanned image by comparison(s) to objects included in a database. Once object recognition program 201 has identified the objects contained within a digital photograph or scanned image, object processing block 202 processes the identified objects to determine the key objects contained within a digital  
5 photograph or scanned image, these key objects are combined into metadata 203 to provide an association between the key objects and a digital photograph or scanned image.

Similarly, as shown in FIGURE 3, a processor or a processing system may accept a user's voice which includes a description of either a scanned image, a textual document, or digital photograph, video, graphics file, audio segment, or other type of data files. As shown  
10 by process 300, translation program 301 preferably converts the received voice into tag information. This tagged information is then processed by semantics processing code 302 which determines the keywords extracted from the spoken data and associated with the scanned document, textual document, or digital photograph. These keywords are then  
15 combined into the metadata in block 303 and provide further information concerning its associated file. Spoken data may be recorded at the time the image was recorded, when it was scanned into the computer or at any other time an association between the image and the spoken words can be established.

FIGURE 4 shows an example of the structure of metadata. Metadata may be any association between the document names or file names and the information contained within  
20 the document (key words and/or key objects) and the voice information (key names) supplied by the user which is also associated with the document or file. The database illustrated in FIGURE 4 shows one example of metadata. In this example, first column 401 consists of the names of the various documents or files contained within the metadata. Columns 402, 403 and 404 preferably contain attributes which describe the files themselves. For example, for  
25 text document 1, two keywords (KEYWORD 1 and KEYWORD 2) were determined through the keyword processing (FIGURE 1, 100) and are associated with text document 1 in columns 402 and 403 respectively. Similarly, the image processing (FIGURE 2, 200) identified two key objects (KEY OBJECT 1 and KEY OBJECT 2) for image 1 and they are associated with image 1 in FIGURE 4. Key names identified through process 300 (FIGURE



3) are also associated with various text documents and image files and are included in column 404. One of ordinary skill in the art would understand that the metadata is not necessarily contained in the database of FIGURE 4, that many representations of the metadata are possible and that FIGURE 4 illustrates only one possible representation. One of ordinary skill in the art would also understand that if a database is used in the implementation of the metadata, the database is not limited to any particular number of columns and rows.

One example of the usefulness of the present invention can be demonstrated by describing how the present invention can be applied to the photographs a typical family takes. Suppose for instance, a family has several hundred photographs. Some of these photographs are in digital format, and others are contained in conventional photographs. The conventional photographs can be scanned into a computer, and each resulting file may be named. The resulting scanned images from the conventional pictures can then, using process 200 of FIGURE 2, undergo the steps of object recognition and object processing where key objects are identified. These key objects can be combined with the image file name to form metadata. Digital photographs can be similarly processed, key objects identified and associated to the file through metadata.

For example, assume ten of the photographs previously mentioned included photographs of various family members playing soccer. In object recognition step 201 (FIGURE 2), these ten photographs of soccer-related events could be identified from objects such as the soccer ball and the soccer goal. Other objects such as grass and trees may also be identified. Object recognition software 201 would identify these various objects within these ten soccer-related pictures. Object recognition software 201 may also identify individuals by their visual characteristics who appeared in the image files. These individuals can be assigned unique identifiers to distinguish them from each other. Once the objects included in the ten soccer-related pictures. Object processing step 202 would determine which objects in the pictures are important and should be kept track of. Object recognition step 201 may have also identified, in addition to the soccer ball and the individuals present in the picture, that the game was played on grass, that the games were played during daylight hours, that there were trees in the background, or a number of other characteristics of the ten soccer-related pictures.

In object processing step 202, process 200 identifies the number of objects which should be included within the metadata associated with this image file. The maximum numbers of objects to be included for each image file in the metadata may be defined by the user, may be included as a default in the processing software, may be obtained from a corresponding table or file format, etcetera. Once object processing step 202 has identified the key objects, the key objects are associated with the image file in the metadata in step 203. Process 200 of FIGURE 2 may be performed at the time the image was scanned, at a later time as defined by the user, or at any other time as defined by the software and/or the user.

Once the ten soccer-scanned photographs are processed by the system, process 300 of FIGURE 3 enables the user to associate additional information with each picture. For instance, referring back to the conventional ten soccer photographs, a first soccer image can be displayed to the user on the screen, and the user can identify the individuals contained within the photographic image, their ages, their relationship to the user, the date and/or time of the soccer game, the circumstances in which the soccer game was played and any other information the user decides to associate with the scanned image. In this example, the user, while viewing the first soccer image may identify two individuals on the soccer field as their son Dominick and their daughter Emily. The user may also indicate that in the photograph Dominick is 6 and Emily is 7, that the soccer game was Dominick's first soccer game and that during this soccer game Emily scored her first goal. This information about the photograph may be provided by text input using a keyboard, designating menu items using a mouse or other positional input device, speech-to-text processing, etc. The information supplied by the user is translated in step 301 of FIGURE 3 into tags that are associated with the scanned image. Semantics processing step 302 may be included, but is not necessary. For instance, if the user simply said "Dominick, Emily, Dominick age 6, Emily age 7, Dominick's first soccer game, Emily's first goal"; the user has identified to the system the keywords the user would like the system to associate with the scanned image. If, however, the user supplies the information to the system in the form of a conversation or a narrative, semantic processing step 302 preferably will be used to extract the key attributes from the narrative. Once the key attributes or key names are identified and associated with the scanned image, this information

is combined into the metadata in step 303. Digital photographs can similarly be associated with key objects and key names.

Once the system has a name associated with an object, this information can be maintained within an associated database so that the object is correctly identified in the future. For instance, in this example, when process 200 of FIGURE 2 was first performed on the first soccer picture, a soccer ball, two individuals, the grass field, daylight and the trees in the background were identified as objects by object recognition step 201. However, at that time, object recognition step 201 was unable to assign unique identifiers to two individuals since object recognition step 201 had no way to associate names with the specific individuals. These identifiers can be used to later associate the individual name with their image. Once the user, using process 300 of FIGURE 3, identifies the two individuals in our example, Dominick and Emily, Dominick and his associated image as well as Emily and her associated image are stored in the object recognition database for future identification. An association between images of Dominick and Emily from other stored images can now be made and previously assigned unique identifiers can be replaced with the individual's name.

Once the keywords, key objects and key names are associated with files the metadata can be used to identify specific files. The metadata now includes in, connection with the soccer picture number information, identification of the soccer ball, Dominick, Emily, the grass field, the trees, Dominick's age at the time of the picture, Emily's age at the time of the picture, the fact that the picture is of Dominick's first game and Emily's first soccer goal, and any other information entered by the user or extracted by the software. The user can now perform searches of the metadata to identify specific pictures from a number of other pictures. For instance if the user queries the system to identify all pictures which are soccer-related, the ten soccer pictures identified previously would be indicated. Additionally, the user can also query the metadata as to when Emily first scored a soccer goal, and the metadata would be able to identify the picture which corresponds to that event.

Image files which began as digital photographs may be similarly processed by process 200 of FIGURE 200 and key names associated with the photograph through process 300 of

FIGURE 3. Similarly, textual files can have key names associated with the textual files as depicted by process 300 of FIGURE 3.

FIGURE 5 is a diagram of an image storage and retrieval system which implements the current invention. In FIGURE 5, imaging device 501, which may include microphone 502, is attached to input/output (I/O) device 503 of processor 504. Processor 504 may be, for example, a document processing engine. Processor 504 is connected to display 505, keyboard 506, preferably microphone 507 and memory 508. Within processor 504, or attached to processor 504, are voice recognition or speech recognition 509 capability, search engine 510 and image recognition capability 511. Imaging device 501 may be a digital camera, a scanner or any other device which allows photographic or image data to be entered into, and processed by processor 504. Microphone 502, if present, may allow a user to record and associate spoken data with a specific image. The imaging data, and any associated spoken data enters processor 504 through I/O device 503. I/O device 503 may also include a disk drive, tape drive, CD, DVD or any other storage device which can be used to introduce image, textual or digital documents or files into processor 504.

Display 505 allows the user to visualize the images, photographs or textual documents as they are associated with keywords, key names or key objects. These associations may be made via user input through keyboard 506, microphone 507 or from image or textual semantics processing 512 capabilities of processor 504. Image recognition 511 capabilities are included in processor 504 for the identification of specific images within image files or photographs. A voice recognition capability translates spoken data received via microphone 502, microphone 508 or I/O device 503 into textual format for inclusion into metadata. Search engine 510 allows the user to process specific metadata information and allows the identification of specific files of interest.

As one of ordinary skill in the art will readily appreciate from the disclosure of the present invention, processes, machines, manufacture, compositions of matter, means, methods, or steps, presently existing or later to be developed that perform substantially the same function or achieve substantially the same result as the corresponding embodiments described herein may be utilized according to the present invention. Accordingly, the

appended claims are intended to include within their scope such processes, machines, manufacture, compositions of matter, means, methods, or steps. Additionally, while a database implementation of the metadata has been described, any searchable association between the file names and the key words, key names and key objects can also be used to implement the metadata.

5

09873687.060404